DOCUMENT RESUME

ED 088 982 UD 014 108

AUTHOR Coleman, James S.
TITLE Effects of School on Learning: The IEA Findings.
REPORT NO IEA-HAR-4
PUB DATE Nov 73
NOTE 47p.; Paper presented at the Conference on Educational Achievement, Harvard Univ., Cambridge, Mass., November 1973

EDRS PRICE MF-$0.75 HC-$1.85
DESCRIPTORS *Academic Achievement; *Cross Cultural Studies; Demography; Elementary School Students; Family Characteristics; *Family Influence; International Education; Multiple Regression Analysis; *Research Methodology; *School Conditions; School Demography; School Statistics; Secondary School Students; Statistical Analysis; Student Characteristics

ABSTRACT

This paper investigates the effects of school on learning as exhibited in this research. In all three studies, there were regressional analyses carried out in which the total variation to be accounted for was that between school averages, and regression analyses in which the total variation to be accounted for was the full variation between individual students. In examining overall effects of school variables, the between-student analysis should be used even though it includes a lot of variance due to individual differences within schools. It appears better to recognize that school variables can account only for that fraction of the total variance that lies between schools, and to use the between-student analysis for the study of effects of school variables. In the use of the between-student analysis, six countries are focused on which engaged in testing in literature, reading, and science: Chile, England, Finland, Italy, Sweden, and the United States. In addition, the analysis is confined to the 10 and 14 year-olds. Several general results are useful to state. For all three subjects, the total effect of home background is considerably greater than the total direct effect of school variables. Comparisons between ages 10 and 14 show a slightly higher total direct effect of schools at age 14 than at age 10 but a smaller proportion of it independent of home background. The data show fairly conclusively that reading achievement is more fully an outgrowth of home influences than are either of the two subjects, less a function of what takes place at school. (Author/JM)

BEST COPY AVAILABLE

ED 088982

IEA/HAR/4

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Effects of School on Learning: The IEA Findings

by

Professor James S. Coleman

University of Chicago
Chicago, Illinois

BEST COPY AVAILABLE

presented at

Conference on Educational Achievement
Harvard University, November 1973

UD 014108

Effects of School on Learning: The IEA Findings

James S. Coleman

University of Chicago

One's first response to the IEA publications on science education, reading comprehension, and literature must be one of amazement and respect: amazement that such a massive set of studies of cross-national achievement of children could be successfully carried out, and respect for those who have done so. It is important to record this first impression, because the studies constitute the best models in existence for cross-national research on social institutions and social behavior. This fact should not be lost in the detailed comparisons, secondary analyses, and critiques of the three studies. With their publication, the comparative study of the functioning of different societies has made an important advance.

This paper, however, is not directed at the most salient results of the research, the differences between countries and the differences between subjects, but at a different question: the effects of school on learning as exhibited in this research. This question could - and has - been asked through research within a single country, on a single subject-matter, and at a single age. But the special virtue of this research is that it covers a number of countries, three age levels, and three subject-matters. For each of the subject-matters separately, the IEA authors have themselves examined this question. Indeed, in all three studies, this question occupies significant portions of the book. Beyond this, however, there are some things to be learned from comparing the separate studies, and I will attempt some such comparisons.

I will divide my comments into two parts, the first methodological and

the second substantive. It is unfortunate that the first is much the larger, for this is an indication that we are still in the early stages of such work, where methodological issues, rather than substantive results, constitute the largest portion of our discourse. Those whose interest is wholly in the substantive results can turn directly to Part II on page 35.

I. Methodological Issues

In all three of these studies, a particular strategy was employed in evaluating the relative importance of different classes of variables. In both the between-school analyses and overall analyses, variables were divided into four "blocks," labelled Blocks 1, 2, 3, and 4. The blocks are roughly defined as follows (the operational definitions differ somewhat from study to study, from between-school to overall analysis, and from Population I to II to IV, but the intent is the same in all cases):

Block 1 Home background, including age and sex of child.

Block 2 Type of school and type of program, for all countries and grade levels in which there was differentiation of program or of school or both.

Block 3 School and Instructional Variables

Block 4 "Kindred" variables, that were not seen as either necessarily prior to achievement nor necessarily consequent upon it, but as possibly either. These were variables such as interest in the subject and motivation.

Interest was centered primarily on variables in Block 3, because they are the variables which can be affected by educational policy and practice. However, interest was also great in Block 2 variables, for somewhat different reasons. Since the type of school a student is in and the type of program he is in (when there are differing school types and differing

programs, as there were in most of the countries) are ordinarily determined on the basis of his performance up to that point, the measure for type of school and type of program shows primarily the degree of differentiation between high- and low-performing students in the system. For example, in the still highly selective secondary system of England, the type of program and type of school accounted for 17 per cent of the total variance among the 14 year olds in science, 13.7 per cent in reading, and 11.9 in literature, the highest among countries in science, second highest in literature, and fourth highest in reading. Thus the interest in Block 2 variables is also related to educational policy, but more nearly to policies of differentiation or selection than to policies designed to directly increase learning. But more of that later.

The interest in Block 1 variables is not quite so direct, except in Population IV, the last year of secondary school. In the lower grades where all students (in the developed countries) are still in school, the variation in family backgrounds reflects the variation in family resources throughout the country, weighted by families' fertility. The size of the effect of this set of variables is thus a product of the variation in family resources (that is, the degree of inequality) in the country, and the transformation of those family resources into (or the effect of that inequality upon) the child's cognitive achievement. Confounded with these two variables is the differential success in measuring the family resources or home backgrounds in different countries. Again, I shall return to this later.

Interest in Block 4, the "kindred" variables, is somewhat less than in the other three blocks for the authors, and will also be of less interest to us here, simply because they are not regarded as wholly causes of cognitive achievement, nor consequences. In further investigations, they may

very well be of interest as dependent variables themselves, because as the literature study shows (p. 284), interest varies widely among countries, and also may well be a function of certain school variables. One such investigation is reported in the literature study (see pp. 410-417). But here, as for the authors themselves, the principal focus will be on cognitive achievement, and while the kindred variables may themselves be important to that achievement, there is nothing in these analyses that would allow such an inference, and they will be, for me, nuisance independent variables, to be disposed of without fanfare.

Now the reason for separating these variables into blocks of this sort was to bring some kind of order into the regression analyses. The problems confronting the analysts were very great, and the number and variety of variables potentially affecting achievement were enormous. Complex analyses prior to the regression analyses themselves were necessary merely to create a reasonably small set of variables without throwing out variables that were important in their effects on achievement. I will not dwell on that elaborate process of data reduction except to commend the analysts. I thought that in Equality of Educational Opportunity our task of analyzing different racial groups and different regions of the country was a massive one; but it is dwarfed by this.

I will go on, however, to be ungracious by commenting on a few difficulties and disorders that still remain. These difficulties can all be summarized by saying that the process was not quite finished. The three studies used, in the end, procedures that were sufficiently different that exact comparisons cannot be made across studies, nor even across the between-school and between-student analyses in the same study. For example, in literature, different variables, from a limited subset, were allowed to enter the analysis

for different countries, based on F-ratio, while in Science, the same variables were used for all countries. Again, in Science the between-school analyses used different weights for different countries in creating composite variables for the between-school analysis, but the same weights over all countries in the between-student analysis. Probably most exasperating to the reader is the fact that the reporting differed quite sharply for the different schools. Some of this merely created inconvenience, as the difference between Reading's use of multiple correlations and Science and Literature's use of squares of multiple correlations, as explained variance. More serious is the fact that non-comparable data are reported, which cannot be recovered: Literature reported (in an appendix) five measures for each individual variable within each block, for the between-student analysis, Reading reported two measures for each individual variable in the between-school analysis, and Science reported no measures for individual variables in either analysis, except for zero-order correlations. And in the between-school analysis, Reading did not separate out, in its reporting, Block I variables from Block 2, so that it is not possible to examine the variation in performance between children in different types of programs or schools, for comparison with the other two studies (see Chapter 7).

But all these are minor quibbles, some caused by nothing more than my difficulties in finding where comparable data were treated in the three studies. I want to discuss serious questions of the methodology used in the regression analyses. For unless it is fully clear what is done through use of this methodology, it will not be possible to draw substantive conclusions about the effect of schools on learning. It will be my contention that the authors were not fully clear about what they were doing, and that this lack of clarity has led them to carry out analyses that prevent one

from answering certain important questions about the effect of schools on learning.

The problem that I want to examine goes to the heart of the procedures used in the regression analyses in all three studies, in both the between-school and between-student analyses. This is the way in which blocks of variables were entered into the analysis. The blocks were used not merely to group variables into sets that were similar in type and interpretation, but also to allow a sequential order in the regression analysis. Block 1 variables, home background and age and sex of child, were entered first, and measures were reported for a regression equation including only them. The measure was explained variance or its square root, the multiple correlation coefficient, and in some analyses, more detailed measures of variables within this block. Block 2 variables, type of program and type of school, were entered second, and measures were reported for them using equations containing Block 1 and Block 2 variables. The measure was the increment to explained variance, a measure obtained by subtracting the explained variance with Block 1 variables alone (R_1^2) from the explained variance with Block 1 and Block 2 variables (R_{12}^2), that is, $R_{12}^2 - R_1^2$. Next was entered Block 3 variables, school variables of numerous sorts, in an equation including Block 1 and Block 2 and Block 3 variables. The measure reported was analogous to that of type of school and program, that is, the increment in variance explained, $R_{123}^2 - R_{12}^2$.

After this, different studies did different things, all entering the kindred variables (Block 4), and Literature going on with Blocks 5 and 6. But these need not concern us.

The major question is an obvious one: what kind of inferences were drawn from the measures reported (R_1^2, $R_{12}^2 - R_1^2$, and $R_{123}^2 - R_{12}^2$), and

what kind can be properly drawn. First, note that the measures are asymmetric for the three blocks. Despite this, inferences were made in each of the studies about the relative effects of home background and school variables, i.e., Blocks 1 and 3 (pages 151-2, 184 in Literature, p. 235, 238, 298-9 in Science, p. 98-100, 121-2 in Reading). In all three studies, the inferences were that home background (Block 1) is more important than school variables (Block 3) and that school variables showed little effect. The terms used to describe school effects varied from study to study, but the characterization of school effects on the basis of these analyses had a certain consistency: "disappointing," (p. 298, Science), "largely negative results" (p. 100, Reading), "there is little the school seems to be able to do to enhance or inhibit" (p. 184, Literature).

To be sure, the quantitative results appear to bear these statements out, and as in earlier studies, the home background variables appear much stronger than the school variables. Furthermore, I am sure that, just as other studies have showed, the home background variables would still have showed great power and the school variables still been "disappointing" if the analysis had been more symmetric. All regression analyses I have seen on these questions, analyzed in whatever way imaginable, would have shown this. But the fact remains that these comparisons are made on the basis of an analysis that is very asymmetric. The measure for the Block 1 variables is R_1^2, while the measure for the Block 3 variables is $R_{123}^2 - R_{12}^2$ (which I shall call in pages following, a_3). It is not the case that R_1^2 is compared with R_3^2, or a_1 ($= R_{123}^2 - R_{23}^2$) with a_3. The fixed order is maintained throughout.

The rationale for this fixed order is best explained in the between-school analysis, and in all three studies a similar rationale was given. In

Science and Literature, a yacht race analogy was used to justify the fixed order, and I will reproduce this justification. Gilbert Peaker is responsible for the yacht race analogy, and one can detect his fine hand in shaping the analyses generally. In my estimation, it is well that this was the case, for there are few if any working analysts of school data for whom I have more respect and admiration than Gilbert Peaker, and I can easily imagine the morass that these analyses might have fallen into if they had not felt his guidance.

Nevertheless, I want to take issue with the use of this asymmetric approach for comparing the relative effects of home background and school variables, and to suggest just what kind of inferences can properly be drawn from the analysis as carried out. The simplest way to do so is by use of the yacht race analogy.

In the attempt to discover effects of school factors on achievement, perhaps the principal villain is the fact that student populations in different schools differ at the outset, and because of this difference, it is not possible merely to judge the quality of a school by the achievements of the students leaving it. It is necessary to control in some way for the variations in student input with which the teachers and staff of the school are confronted. In some way, it is the <u>increment</u> in achievement that the school provides, which should be the measure of the school's quality. If we had good measures of that increment, as well as good measures of the level of various school resources in the same school, it would be possible to establish a relation between the size of the increment and the level of certain resources, and thus to determine which school resources were most important to learning.

The problem lies in establishing the appropriate baseline so that some

estimate of the achievement increment can be made, and most cross-sectional studies have, like this one, attempted to use factors in the student's own background or possibly in the community which can provide an estimate of the student input to the school and thus allow an estimate of the increment of achievement.

The yacht race analogy is this: in a yacht race, in order to give all boat crews a chance that is independent of the size of the boat, size of sails, and other dimensions of the boat, a formula is used that gives each yacht a handicap, based on the expected or average performance of yachts with those dimensions. This places all boats in an equal starting position, so to speak, and the crews have equal chances to win, even if their boats sails are small.

Similarly, thinking of the children's performance in analogy to the yacht's performance, we must recognize that because of the effect of background and other factors, different children have different expected performance apart from their school. Then if we are to obtain a measure of what the school does (in analogy to the crew) to increase this performance, we must give each school a handicap score which is based on the expected performance of students like those in this school.

On these grounds, a school handicap variable was constructed, based primarily on the family background characteristics of students in the school. In addition to the school handicap score, other variables which differed from school to school, such as sex of students and age, were included in Block 1 because of the different expected or average performance of students of different sex and different age.*

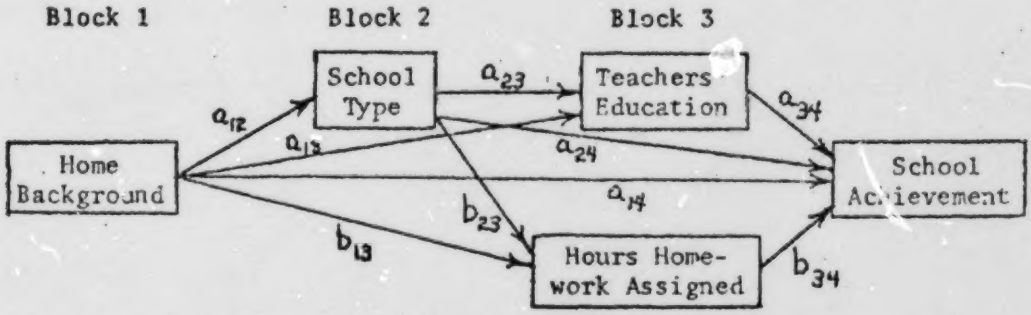
* Age in years was held constant in the analysis; but age in months within that year showed some positive relation to performance in most of the analyses. Boys performed much better in science and girls much better in literature. In these two subjects, sex of the student was one of the most important var-

The analogy appears a good one. But there is one problem, if we are to start the schools out on a completely equal footing for the analysis of Block 2 and Block 3 variables. This is, how do we determine the size of the handicap? In handicapping yachts, the size of the handicap is determined by racing times of yachts of given dimensions with a variety of different crews, to determine the average or expected performance for a yacht of those dimensions. Suppose, however, yachts with larger sails tended to be manned by better crews, so that the performance of those yachts, averaged over many crews, included an increment of performance due to the better crews, rather than only an increment due to the larger sails. Then the handicap score will be wrong, because of the correlation of good crews with larger-sailed boats. The average performance of larger-sailed boats will be overestimated, and the subsequent performance of crews sailing those boats will be underestimated. It is only because a handicap score for boats with given dimensions can be made independent of the quality of the crews that the handicapping works correctly.

But the hypothetical possibilities I have described for the yacht handicapping is just what exists in the schools. There <u>is</u> a correlation between student input, as approximated by home background variables, and school resources, such as teacher quality and the like. Consequently, to develop a handicap for the schools by first accounting for all the variance possible through home background variables (entering Block 1 first) extracts out not merely the variance due to student input, but also that due to the school variables that are correlated with student input. Then to compare the effect of school variables with that of home background variables is not appropriate.

iables in the analysis. (In Reading, no separate report was made for sex except in the zero-order correlations, which showed small correlations, favoring girls in 33 countries and boys in 14.)

I am not an advocate of path analysis, but I can use a path diagram similar in principle to that presented for Japan (p. 280-284, Science) to illustrate the point. Assume a home background variable (Block 1), a school type variable (Block 2), and two school variables (Block 3).



The diagram indicates the causal reasoning behind the sequence of blocks, and I have no quarrel whatsoever with this set of a priori causal assumptions. All analyses of school effects that I know suggest this kind of scheme, in which home background precedes, in a causal sequence, both the type of school and particular variables of the school, and in which the type of school (in a differentiated system) precedes the particular school variables.

An alaysis to identify the relative sizes of the causal flows labelled by a_{ij} and b_{ij} in the diagram would require four regression equations: a) school type as dependent variable and home background as independent; b) Teachers' education as dependent, school type and home background as independent; c) hours homework as dependent, school type and home background as independent; and finally, d) school achievement as dependent, and all four others as independent. These do not, with the exception of (d), correspond to the equations used in the analyses under discussion. But no matter. What we want to see is just what the analyses under discussion did correspond to.

1) First, an equation with Block 1 as independent and school achievement as

dependent shows the total effect of the Block 1 variable, through these paths: direct: a_{14}; 2-step: $a_{12}a_{24}$, $a_{13}a_{34}$, and $b_{13}b_{34}$; and 3-step: $a_{12}a_{23}a_{34}$ and $a_{12}b_{23}b_{34}$. The variance explained by Block 1 is explained through these six paths.

2) Second, the equation with Block 1 and Block 2 shows the effect of three paths from equation 1 (a_{14}, $a_{13}a_{34}$, and $b_{13}b_{34}$) plus those from school type, both direct and 2-step: direct: a_{24}; 2-step: $a_{23}a_{34}$ and $b_{23}b_{34}$. When the variance accounted for in equation 1 is subtracted out, what is left is all due to school type, that is, a_{24}, $a_{23}a_{34}$, and $b_{23}b_{34}$: that due to a_{14}, to $a_{13}a_{34}$, and to $b_{13}b_{34}$ is subtracted out. But also subtracted out is a portion of the variance that operated through school type: $a_{12}a_{24}$, $a_{12}a_{23}a_{34}$, and $a_{12}b_{23}b_{34}$. These paths are much less strong than a_{24}, $a_{23}a_{34}$ and $b_{23}b_{34}$ whenever a_{12} is itself small; and it usually is not very large. Nevertheless, some portion of the effect of a_{24}, $a_{23}a_{34}$, and $b_{23}b_{34}$ is subtracted out, the portion depending on the size of a_{12}.

3) The third equation includes Block 1, 2, and 3 variables, and includes variance from two paths in equation 2 (a_{14} and a_{24}) plus that from Block 3 variables: a_{34} and b_{34}. When the variance accounted for in equation 2 is subtracted out, what is left is all due to Block 3, that is a_{34} and b_{34}: that due to $a_{14}$and a_{24} is subtracted out. But also subtracted out is a portion of the variance due to a_{34} and b_{34}. Variance due to all these paths is subtracted out: $a_{13}a_{34}$, $b_{13}b_{34}$, $a_{23}a_{34}$, and $b_{23}b_{34}$. Thus what is left after equation 2 variance is subtracted out is var (a_{34}) - var ($a_{13}a_{34}$) - var ($a_{23}a_{34}$) + var (b_{34}) - var ($b_{13}b_{34}$) - var ($b_{23}b_{34}$). (I am using var ( ) to mean the variance due to a given causal relation.)

Now, let us summarize what variance is contained in the numbers reported for Blocks 1, 2, and 3 in these studies:

Block 1	Block 2	Block 3
R_1^2	$R_{12}^2 - R_1^2$	$R_{123}^2 - R_{12}^2$
var (a_{14})		
var $(a_{12}a_{24})$	var (a_{24}) - var $(a_{12}a_{24})$	
var $(a_{13}a_{34})$		var (a_{34}) - var $(a_{13}a_{34})$
var $(b_{13}b_{34})$		var (b_{34}) - var $(b_{13}b_{34})$
var $(a_{12}a_{23}a_{34})$	var $(a_{23}a_{34})$ - var $(a_{12}a_{23}a_{34})$	- var $(a_{23}a_{34})$
var $(a_{12}b_{23}b_{34})$	var $(b_{23}b_{34})$ - var $(a_{12}b_{23}b_{34})$	- var $(b_{23}b_{34})$

This chart shows the asymmetry introduced by the procedure used in these studies. Note that asymmetry exists in two ways: the variance for Block 1 includes all the variance due both to the direct path and all indirect paths; the variance for Block 3 not only is limited to that from direct paths, but excludes that due to indirect paths from earlier steps.

Now there is nothing wrong with such asymmetry, but it is important to be aware of its implications. Two of its important implications are these: 1. It is not possible to make a comparison of the amount of variance accounted for in different blocks, and say, for example, that Block 1, home background, accounts for much more variance than Block 3. Thus the statements in these studies about the small effects of school variables compared to the effects of home background, which "account for much of the variance," (p. 152, Literature), are of "decided importance," (p. 100, Reading), or are "quite considerable," (p. 235, Science), are not appropriate statements on the basis of the analyses reported. The statements are very likely true, as inspection of the zero-order correlations, the path analysis for Science in Japan, or any of numerous other indicators suggest, but in order to make an explicit comparison, a symmetric analysis would be necessary - for

example, a commonality analysis of the variance R_{123}^2, showing both the unique variance and total variance accounted for by each block in that equation. As it stands, the variance estimates at Block 1 are of total variance due to Block 1 variables, while those at Block 3 are of the unique variance due to Block 3 variables, with those at Block 2 being somewhere in between.

But it is perhaps not even appropriate to want to make such symmetric comparisons between "effect of home" and "effect of school" if one has in mind the kind of causal diagram as shown above, because the home and school occupy quite different places in it: home variables partly determine school variables, but not the other way around. What one might want to specify for Blocks 1 and 3 are these effects:

1) the total effect of variations in Block 1, (home background) variables, both through its impact on school variables and independent of those variables.

2) The total direct effect of variations in Block 3 (school) variables, whether this effect is merely implementing the force of home background (paths $a_{12}a_{23}a_{34}$, $a_{13}a_{34}$, $a_{12}b_{23}b_{34}$, $b_{13}b_{34}$) or independent of it. This may be thought of as the potential direct effect of school variables if they were distributed independently of home background.

3) The direct effect of variations in Block 3 (school) variables as distributed independent of or over and above the force of home background.

4) The direct effect of variations in Block 1 (home background) variables, apart from the force of home background in shaping or selecting schools.

Perhaps other effects are desired as well, but these appear to me the most

important. One might then want a) to compare effects 2 and 4 to compare the direct effect of school with the direct effect of home independent of school. But if one did so he should realize that this does not express the total effect of home, because the home acts to determine the school variable themselves. Consequently, one might want b) to compare 1 and 2, the total effect of variations in home back-round with the total direct effect of school. One might also want c) to compare effects 1 and 4, to determine the proportion of the home's effect that takes place through shaping or selecting the school the child goes to, and the proportion that takes place directly through its impact on the child himself.* And one might want to d) compare effects 2 and 3, to examine what proportion of variations in the school's impact is distributed independently of differences in family background, compared to the proportion that merely reinforces those differences. In addition, one might want e) to compare effects 1 and 3, recognizing that these are total effects of home background variations, but only partial effects of school variations. The idea here might be to compare the total effects of family variations on achievement with the effect on the child's achievement of other variations in society that act through the schools but independent of the home.** Finally, one of the most important comparisons one might want to make lie within a block: for example, one might want f) to compare the relative sizes of the total direct effects of different Block 3 variables

* When Block 2 is included in the consideration, for differentiated school systems further possibilities exist, some of which may be quite important: for example, the effect of home (through whatever means, including direct effects on the child's performance) in determining the type of school attended vs. the home's effect apart from this; the effect of the type of school in itself on achievement, apart from the specific resources that exist in that type of school (for example, through the selected student body) vs. the effect of type of school through the school resources it provides.

** This is not the same as comparing the total effect of the family on the child's achievement with the total effect of the other aspects of society through school on his achievement. Since we are not testing the absence <u>vs.</u>

(comparisons within (2) above), or comparison of the relative sizes of the effects of different Block 3 variables that, as distributed, are independent of the force of home background (comparisons within (3) above).

That is, given a set of a priori assumptions about causal flows, the kind of comparisons one wants to make may very well not be symmetric, and it is not reasonable only to think of comparing "home background effects" and "school effects" without further specifications.

Now given the kind of analyses that have in fact been carried out in these studies, which of these comparisons is possible? What has been measured through the sequential introduction of blocks, and the reporting of R_1^2, $R_{12}^2 - R_1^2$, and $R_{123}^2 - R_{12}^2$ are effects (1) and (3) listed above. Consequently, it is possible, among countries to show two things, if we first consider blocks as wholes and considering only Blocks 1 and 3, neglecting Block 2 for the present, or combining it with 1: the total effect of home background, independently of whether it operates through shaping the school or shaping the child, and the effect of school variables independent of the force of home background. Of the various comparisons, (a) through (f) above, which of them are possible with these data? Only two, (e) and (f).

Now this is not disastrous in any way. If the analyses had been carried out so that all comparisons (a) through (f) could be made, then there would have been far more to do than possible within the confines of these volumes. But it is important to recognize what kind of comparisons the data do allow,

the presence of the family or the absence vs. the presence of other societal forces, all that is possible is testing the total effect of variation within families vs. the effect of variation within schools induced by other societal forces. An example of a case in which the other societal forces were almost wholly absent was in Prince Edward County, Virginia, several years ago, when the public schools were closed for several years. What happened then was an intensification of the effect of home background. Families with any financial resources to spare, cognitive skillsof their own to transmit to their children, or interest in their children's education used their money to create private schools for their children or to

so that the inferences drawn are not incorrect.

Everything I have said so far has been a necessary methodological preface to any discussion of substantive results about effects of schools on learning as shown by these studies. For until it is clear what inferences are possible with these data and what are not, then it is not reasonable to begin to examine the substantive results. But before beginning this examination, it is necessary to carry further the methodological questions, asking about the use of "variance added" as a measure of the effect of a class of variables.

The Use of "Variance Added"

In all these studies, "explained variance added" was used as the principal measure of the effect of a class of variables, although in two of the studies an additional measure was used: standardized regression weights in the reading study, and unique contribution to explained variance in the literature study. The use of this measure is dictated in part by the fact that it is one of the few easily obtainable measures of the effect of a set of variables, as distinct from the effect of a single variable, for which regression coefficients and standardized regression coefficients are available.

The limitation of this measure, in terms of what it tells about the effects of the block of school variables, were discussed above. But even

send their children to private schools, or tutored them at home, while families with little money or interest in their children's education, and little capability of teaching them (mostly black) neither sent their children to school nor tutored them. The results were greatly increased differences in cognitive skills as a function of family background - indicating through their absence the effect that uniform societal forces outside the family have in reducing the effect of family variations on achievement. But in the present studies, the only children tested were children in school, and except at the highest age level, nearly all the children in the population were in school, except in Chile (age 10, 94%,

accepting the use of a measure to show the effect of a block of variables as distributed, over and above the effect of Block 1 and 2 variables, the question is, is it the correct measure? I think not.

"Explained variance" (R^2) is the square of the multiple correlation coefficient (R) and R^2 has come to be fashionable as a measure of the power of the set of independent variables in explaining variation in the dependent variable. "Explained variance added" by a set of variables then becomes the difference between R^2 of the regression with these variables and R^2 of the regression without them. My point is that all this is mistaken, and instead that the multiple correlation coefficient itself should be used, that is, R instead of R^2, and the "correlation added," or "explained variation added," rather than "explained variance added" be used as a measure of the contribution of a set of variables over and above the effect of another set.

In arguing for measures other than the added explained variance or unique variance, I must point out that I have until very recently used these measures in all my work, and in Equality of Educational Opportunity, we used them. Although I still believe these measures are superior for the present purpose to others currently in use, I think they are not as good as the alternatives I am proposing here.

The basis for this argument is rather straightforward. Consider a set of three regression equations such as the ones below for England in Literature (p. 163). The variance explained by a given set of variables, s, is denoted by R_s^2, and the variance explained by that set excluding Block i variables is denoted by R_{s-i}^2.

age 14, 71%0, India (age 10, 50%, age 14, 25%), Iran (age 10, 75%, age 14, 25%), Italy (age 14, 55%), (Science, p. 57). Consequently, such total societal forces could not be measured.

i	R_s^2 (variance explained)	R_s (mult. correl.)	$R_s^2 - R_{s-i}^2$ (variance added)	$R_s - R_{s-i}$ (variation added)
1. Block 1 (home background)	.252	.502	.252	.502
2. Blocks 1 and 2 (+ type of school & program)	.371	.609	.119	.107
3. Blocks 1, 2, 3 (+ instructional variables)	.410	.640	.039	.031

The measures used in the studies (R_s^2 and $R_s^2 - R_{s-i}^2$) are reported, along with the measures I propose, R_s and $R_s - R_{s-i}$. Now suppose, looking first at the first equation, we consider an attempt to explain cognitive achievement in literature by a single composite variable, "home background." That composite variable is a linear combination of the variables that go into it which minimizes the sum of squared deviations of the predicted dependent variable from the actual one. That is, it is a composite variable which is a weighted sum of the various home background variables, the weights being the regression coefficients themselves. Now since we conceive of such a variable, "home background," we can ask what would be the expected performance of a child at a given percentile of home background - say the 25th and 75th percentiles? What is meant by the "25th percentile of home background" is a family background low in those resources that contribute to a child's performance in literature. The particular combination of resources might differ: one child with a family background at the 25th percentile might have low father's education and high mother's education, while another might have high father's education and low mother's education. The two backgrounds are equivalent only in that on the average, homes with these two combinations of resources both produce children at the 25th percentile of the distribution of predicted performance.

Now the answer to the question of what would be the expected perfor-

mance of a child at the 25th and 75th percentiles of family background is not, as one might at first glance expect, the 25th and 75th percentiles of actual performance. It is, rather, the 25th and 75th percentiles of predicted performance, which has a smaller standard deviation than the distribution of actual performance. The ratio of the standard deviation of predicted performance to actual performance is simply R_1, the multiple correlation between achievement and the home background variables in Block 1. Thus if the average performance is 50, and the standard deviation is 10, then assuming a normal distribution, a child at the 25th percentile will score 43.26 while a child at the 75th percentile will score 56.74, that is 6.74 points below and above the average, respectively. A child in England at the 25th percentile of home background will score at $50 - 6.74\ R_1$ or 46.61, while a child at the 75th percentile of home background will score at $50 + 6.74\ R_1$, or 53.39. That is, if we know that the score for a given percentile is a certain distance below or above the mean, then the score for that same percentile on this composite variable will be R_1 times that distance - in this case, 50.2 percent of that distance. Thus we can reasonably say that this composite variable, home background, accounts for 50.2 percent of the variation (not variance) in achievement. What is true at the 25th percentile is true at all percentiles - the child at a given percentile on home background is R_1 times the distance from the mean in achievement that the child at the same percentile in performance is. Or to look at it differently, if we know the difference in achievement score between any two percentiles (a difference of 13.48 points for the 25th and 75th in the example above), then the difference in achievement score of the average child with a home background at those same two percentiles will be R_1 times that difference, or $13.48\ (.502) = 6.78$ in this case.

It is useful also to recognize that R_1 is the standardized regression coefficient β_1 of the composite variable home background when no other variables are controlled (i.e., including effects of all variables through which it is correlated). Ordinarily, one does not use standardized regression coefficients in single-variable regression analysis, however, this identity between R_1 and β_1 will be important to the later discussion.

By extension, the same argument given above now holds if we create a composite variable from home background and type of program and school, that is, from the Block 1 and Block 2 variables combined. We can ask the question about two children at the 75th and 25th percentile of a new compound variable, made up of home background and type of program and school: what proportion of the total distance above the mean to the 75th percentile is the child who is at the 75th percentile in this combination of home and school resources? That proportion is R_{12}, where R_{12} is the multiple correlation of achievement with home background and type of program and school. R_{12} again is a standardized regression coefficient β_{12}, for composite variable made up of Blocks 1 and 2, when no other variables are controlled. Of course it may not make much sense to define a new variable as "home and school resources" (where we mean by school resources type of school and type of program). If not, then the usefulness of R_{12} is not very great. But in any case, our interest is not so much in the explanatory power of the best combination of home and school resources as in the explanatory power of school resources themselves, apart from home resources. The natural thing to do then seems to be to subtract out the variation that can be explained by home alone, and consider only the additional variation that is accounted for by the Block 2 variables, that is, $R_{12} - R_1$ or thinking in terms of standardized regression coefficients, $\beta_{12} - \beta_1$. But now let us see

what this tells us. It tells us the additional explanatory power of the best compound variable made up of home resources and school type and program, beyond the explanatory power of the best compound variable of home resources by themselves. What this tells is the additional power to explain achievement brought in through type of school and program that is distributed independently of home background, that is, effect of type (3) discussed on page 14 above, the kind of effect that $R_{12}^2 - R_1^2$ measures imperfectly. But it does not tell us in any way the total direct effect of those Block 2 variables, that is, the effect of type (2) on page 14 above. What would such an effect be? We can think of an hypothetical experiment something as follows: suppose for two children, the Block 1 variables, home background, were at their average position for the population. Then if for one of the children, the Block 2 variables were held at the population average, his predicted score would be exactly at the population average. But for the other child, the Block 2 variables are put at their 75th percentile level. What we mean by "their 75th percentile level" is this: in this equation, including the Block 1 variables, we find the linear combination of Block 2 variables that in the presence of Block 1 variables has the highest partial correlation with achievement. This becomes for us a new variable which we labeled "type of program and school," or "Block 2". Then we find the 75th percentile position on this new variable (which we can assume is normally distributed, for purposes of the hypothetical experiment and the measure that this experiment is leading toward). And the second child is at that 75th percentile level, along with his 50th percentile family background level. Our question then becomes, what is the predicted achievement of this child at the 50th percentile of family background (Block 1) and the 75th percentile of Block 2? Further,

thinking of the difference between the average score (50th percentile of Block 1 and 50th percentile of Block 2) and the 75th percentile score (which is simply .6745 times the standard deviation of achievement), what proportion of that distance is covered by the score of our second child? Whatever measure will give us the answer to that question is the measure of the total direct effect of the type of program and school compound variable.

We can easily see how, if we truly went to the trouble of making up that compound variable as described above, such a measure would be directly forthcoming from the regression equation. We can see how by directly calculating the three scores in question:

1) 75th percentile score: $\bar{y} + .674\sigma_y$
2) 50th percentile or average score
   (with average Block 1 and Block 2 variables) $\bar{y}$
3) predicted score of a child
   with average Block 1 variables (x_1) and 75th percentile on Block 2 variable (x_2)

If the regression equation is $y = a + b_{1.2}x_1 + b_{2.1}x_2$, then predicted score y^*,

$$y^* = a + b_{1.2}\bar{x}_1 + b_{2.1}(\bar{x}_2 + .674\sigma_{x_2})$$
$$= a + b_{1.2}\bar{x}_1 + b_{2.1}\bar{x}_2 + b_{2.1}(.674\sigma_{x_2})$$
$$= \bar{y} + b_{2.1}(.674\sigma_{x_2}).$$

But if $b_{2.1}$ is the raw regression coefficient of Block 2 variables, controlling on Block 1, and $\beta_{2.1}$ the standardized one, then $b_{2.1}\, x_2 = \beta_{2.1}\sigma_y$. Thus

$$y^* = \bar{y} + \beta_{2.1}(.674\sigma_y)$$

Therefore, the proportion of the distance between scores (1) and (2) covered by (3) is:

$$\frac{\bar{y} + \beta_{2.1}(.674\sigma_y) - \bar{y}}{\bar{y} + .674\sigma_y - \bar{y}} = \beta_{2.1}$$

Thus the desired measure, showing the proportin of variation in achievement that the Block 2 variable will explain when Block 1 variables are held fixed is merely the standardized regression coefficient in the equation containing Block 1 and Block 2.

The approach taken here toward carrying out hypothetical experiments with blocks of variables can be carried further. Some examples will show the generality of this approach toward measures of effects in regression analysis, and will make evident that it should not be practiced merely through blind calculation of standardized regression coefficients without regard for precisely what is the desired measure. One hypothetical experiment would be to ask the achievement between a child who is at the 25th percentile in home background (Block 1), and at the 75th percentile in school type and program(Block 2). Is this score above or below the mean, and what proportion of the distance from the mean to the 75th or 25th percentile is it? The score is:

$$\begin{aligned} y^* &= a + b_{1.2}(\bar{x}_1 - .674\sigma_{x_1}) + b_{2.1}(\bar{x}_2 + .674\sigma_{x_2}) \\ &= a + b_{1.2}\bar{x}_1 + b_{2.1}\bar{x}_2 - b_{1.2}\cdot.674\sigma_{x_1} + b_{2.1}\cdot.674\sigma_{x_2} \\ &= \bar{y} - b_{1.2}\cdot.674\sigma_{x_1} + b_{2.1}\cdot.674\sigma_{x_2}. \end{aligned}$$

But $b_{1.2}\sigma_{x_1} = \beta_{1.2}\sigma_y$, and $b_{2.1}\sigma_{x_2} = \beta_{2.1}\sigma_y$, so that

$$y^* = \bar{y} + .674\sigma_y(\beta_{2.1} - \beta_{1.2}).$$

If the quantity in parentheses is positive, the score is above the mean. The score at the 75th percentile is $y + .674\sigma_y$, so that the proportion of distance from $\bar{y}$ to the 75th percentile is

$$\frac{\bar{y} + .674\sigma_y(\beta_{2.1} - \beta_{1.2}) - \bar{y}}{\bar{y} + .674\sigma_y - \bar{y}} = \beta_{2.1} - \beta_{1.2}$$

If the quantity is below the mean, the proportion of the distance to the 25th percentile would be calculated in the same way, but would be $\beta_{1.2} - \beta_{2.1}$.

A second hypothetical experiment would be this: suppose a child is one standard deviation below the mean in home background (Block 1) resources, and at the mean in school type (Block 2). Then how many standard deviations above the mean should his school resources (Block 3) be to make his predicted score be at the mean? This is answered be setting up the appropriate equation with the desired number of standard deviations represented by an unknown, α.

$$\bar{y} = a + b_{1.23}(\bar{x}_1 - \sigma_{x_1}) + b_{2.13}\bar{x}_2 + b_{3.12}(\bar{x}_3 + \alpha\sigma_{x_2})$$

$$= a + b_{1.23}\bar{x}_1 + b_{2.13}\bar{x}_2 + b_{3.12}\bar{x}_3 - b_{1.23}\sigma_{x_1} + \alpha b_{3.12}\sigma_{x_2}$$

$$\bar{y} = \bar{y} - b_{1.23}\sigma_{x_1} + \alpha b_{3.12}\sigma_{x_2}$$

$$0 = -\beta_{1.23}\sigma_y + \alpha\beta_{3.12}\sigma_y$$

$$\alpha = \frac{\beta_{1.23}}{\beta_{3.12}}$$

These hypothetical experiments show the generality of maniplulations among standardized regression coefficients as a way of measuring the effects of changing various independent variables by some standard distance related to their distribution in society, such as one standard deviation.

It is important to recognize that even if we had the standardized regression coefficients, they would be no panacea. It tells us only the effect of type (2) on page 14. Suppose it is empirically the case that children in the high-performance schools and programs (Block 2) are nearly always from good home backgrounds (Block 1). Then it is quite artificial, in one sense, to perform the hypothetical experiment discussed above, for seldom does a child from an average home background attend a 75th percentile school. To know the potential for higher performance of a child in a selective secondary school with average home background is of purely academic interest if it infrequently happens. What is of greater interest

is the amount of variation that the schools in fact account for independently of the family, that is, $R_{12} - R_1$, or $\beta_{12} - \beta_1$. But if there is not such a high correspondence between Block 1 and Block 2, the potential effect is in fact realized for some children, and it is of more than academic interest.*

In the above discussion, it was assumed that Block 2 consisted of a single compound variable. But if a block consists of a number of variables, how do we get a measure for the overall impact of the block of variables, considered conceptually as a compound variable (e.g., "school variables," "instructional variables," or "learning conditions," as the authors termed Block 3 variables), but in practice left as a set of variables? One of the appealing aspects of "added explained variance" ($R_s^2 - R_{s-i}^2$) or "added explained variation", ($R_s - R_{s-i}$) is that it serves as a global measure for the whole set of variables that are entered in a given block. Since the blocks have a coherent meaning or interpretation it is useful to have a measure, comparable to the standardized regression coefficient, for the block considered as a compound variable, even though it is a number of separate variables. This measure would tell us the proportion of variation that Block 2 variables, considered as the best-predicting compound, will explain when all Block 1 variables are held fixed. But it turns out this is not quite so simple to obtain as the added explained variance, $R_{12}^2 - R_1^2$, or the added explained variation, $R_{12} - R_1$. Conceptually, there is no problem. In fact, one way of defining

* I have discussed Block 2 variables as if they were school variables that had an effect on learning. In fact, as the authors point out, they are primarily indicators of differential selection of differently-performing students into different schools, and thus surrogates for unmeasured variations in student input into the schools. They show not how much learning takes place in the schools, but how much selection.

$\beta_{2.1}$ gives an interesting insight into the conceptual difference between the "added variance" and this measure. The added variance for Block 2 is $R_{12}^2 - R_1^2$, and the square of the standardized regression coefficient $\beta_{2.1}$ is this same quantity divided by one minus the square of the correlation between the Block 2 compound and the Block 1 compound, which we will label R_{21}. (In the equation below, we will introduce the subscript 4 for the dependent variable of achievement, to prevent confusion.) The equation for $\beta_{42.1}^2$ is:

$$\beta_{42.1}^2 = \frac{R_{4.12}^2 - R_{41}^2}{1 - R_{21}^2}$$

The division by $1 - R_{21}^2$, which is the only difference between β^2 and added variance, shows the conceptual difference between the two. For R_{21}^2 is the proportion of variance in the Block 2 variable compound that is accounted for by the compound of Block 1 variable (not a compound designed to best explain the Block 2 compound, but the compound based in regression weights in the equation with achievement as dependent variable, and Blocks 1 and 2 as independent). Thus the added variance is merely the square of the regression coefficient discounted by the variance that is common with the Block 1 compound.

To calculate this standardized regression coefficient, it would be possible to first carry out the full regression of achievement on Block 1 and 2 variables, then to create a new compound variable from the Block 2 variables, by using the regression coefficients as weights, recalculate the correlation matrices including the newly-defined variable, and then carry out another regression analyses, in which the new compound Block 2 variable replaces the set of variables from which it was compounded. The new regression equation is identical to the preceding one, in total variance

explained and in regression coefficients for the unaffected Block 1 variables. For the Block 2 compound, the raw regression coefficient and the standardized coefficient are identical, and are the desired measure: the proportion of variation in achievement that will be accounted for by Block 2 variables when Block 1 variables are held fixed.

This kind of recalculation is, however, technically cumbersome. It involves recalculation of the covariance matrix every time a new compound is created. Instead of that, it is possible to proceed in any of several ways that make it possible to calculate these standardized regression coefficients without recalculating the correlation matrices. I will discuss two such ways in the Appendix 2, but here it is sufficient to note that if we have the multiple correlation between the Block 2 variables alone and achievement, as well as the multiple correlation between Block 1 and achievement and between Blocks 1 and 2 and achievement, it is possible to calculate the standardized coefficient directly. For in the equation given earlier, only $R_{4.12}^2$, R_{41}^2, and R_{21}^2 are necessary to calculate $\beta_{42.1}$. $R_{4.12}$ and R_{41} are multiple correlations of 1 and 2 with 4 and 1 with 4, respectively. R_{21} is not easy to obtain directly, but if R_{42}, which is easy to obtain, is known along with $R_{4.21}$ and R_{41}, then R_{21} may be calculated from them, as described in the appendix. The three multiple correlations necessary to calculate R_{21} are not given in these studies, but two of them are, and it is possible to calculate reasonable upper and lower bounds on the desired standardized regression coefficient, by the method described in the appendix 2. Thus in these studies, we can approximate the desired standardized regression coefficient $\beta_{2.1}$ for Block 2 variables by

$$\beta_{2.1} \approx \frac{\sqrt{R_{12}^2 - R_1^2}}{2} \left(1 + \frac{1}{\sqrt{1 - R_1^2}}\right)$$

and for Block 3 variables

$$\beta_{3.12} \approx \frac{\sqrt{R_{123}^2 - R_{12}^2}}{2} \left(1 + \frac{1}{\sqrt{1 - R_{12}^2}}\right)$$

This will give us estimates of the total direct effect of Block 2 and 3 variables to complement the direct effects distributed independently. It is interesting that although this measure is very different from the added explained variance, it can be roughly estimated from the same two quantities, R_s^2 and R_{s-i}^2, which are used to calculate the added explained variance.

Now it becomes possible to indicate what would be the approximate measures for the four kinds of effects discussed for Blocks 1 and 3 earlier in the paper. The notations used will be subscript i for Block i, with subscript 4 denoting the dependent variable, achievement; R_{4i} as the multiple correlation of Block i alone with achievement, $R_{4.ijk}$ as the multiple correlation of Blocks i,j,k with achievement; $_{4i.jk}$ as the standardized multiple regression coefficient of achievement on the Block i variables, considered as a compound, when Blocks j and k are controlled.

1. Total effect of variations in Block 1 variables, both through through school and outside it: R_{41} $(=\beta_{41})$
2. Direct effect of Block 3 variables, whether this effect implements the force of home and school type through its distribution, or not: $\beta_{43.21}$
3. Independent direct effect of Block 3 variables as distributed, in implementing societal variations beyond the home: $R_{4.321} - R_{4.21}$ $(=\beta_{4.321} - \beta_{4.21})$
4. Direct effect of Block 1 variables, apart from their effects through shaping or selecting schools: $\beta_{41.23}$

There is one important point that should be made about the virtues of the "added explained variance" or "unique variance" measures used in these studies. Although they are not estimates, as are the measures I have proposed here, of the amount of variation explained by particular variables or blocks of variables, they have some special virtues. This is best seen by writing again the equation for $\beta_{43.12}$:

$$\beta_{43.12} = \frac{\sqrt{R_{4.123}^2 - R_{4.12}^2}}{\sqrt{1 - R_{3.12}^2}}$$

Now the added, or unique variance is simply $\beta_{43.12}^2 (1 - R_{3.12}^2)$, that is, the square of the standardized regression coefficient, the measure I have proposed, times the variance in Block 3 (or variable 3) which is not accounted for by the other independent variables, 1 and 2. The crucial difference between the two is not the square vs. non-square, but the discounting by $1 - R_{3.12}^2$ or not. The regression coefficient β in effect says to take the (square root of) the added explained variance, but to take into account that a portion of the variance of variable 3, $R_{3.12}^2$, had no chance to be effective, because that variance coincided with the variance of variables 1 and 2 which already accounted for variance $R_{4.12}^2$. Thus it multiplies the (square root of) the added explained variance by a ratio $1/(\sqrt{1 - R_{3.12}^2})$, which inflates $\sqrt{R_{4.123}^2 - R_{4.12}^2}$ back up to what its estimated value would be if variables 1 and 2 were not correlated with it.

The only difficulty with this procedure is that if $R_{3.12}^2$ is high, then this constitutes a great inflation, and in effect a large extrapolation of the variation it has explained - an extrapolation that could be mistaken.

For this reason, added explained variance, $R_{4.123}^2 - R_{4.12}^2$ (or what is

called in Literature b^2/c for single variables) is a valuable conservative statement about the effect of variable or Block 3. It is particularly valuable for estimating the relative effects of different school variables which have different correlations with home background variables, and for which the extrapolation due to $R_{3.12}^2$ might be very different, and possibly a bad extrapolation to use. It is used in this way in these studies (particularly Literature) for the examination of individual variables. However, its depressed value due to $R_{3.12}^2$, and due to the use of the squared rather than the square root form, leads to an unwarranted pessimism about the size of the effects.

I would prefer that the square root be used, because that would have an explicit meaning in terms of variation of the dependent variable analogous to those described earlier. The meaning would be this: $\sqrt{R_{4.123}^2 - R_{4.12}^2}$, or equivalently $\beta_{43.12}\left(\sqrt{1 - R_{3.12}^2}\right)$, is the proportion of the distance between scores of the student at percentile A on achievement and the score of the student at percentile B which would be covered by changing the Block 3 composite variable for a student from whatever percentile it was at when the Block 1 and 2 composite was at percentile A, to percentile B. (If there is a strong correlation between Block 3 and Blocks 1 and 2, the original percentile of the Block 3 composite will already be some distance to percentile B, because of the correlation. If not, $R_{3.12}^2$ close to zero, then the Block 3 composite will originally be close to percentile A.)

Thus the use of added explained variance, but in its square root form, consitutes a useful statement about the effect of Block 3 variables as distributed. There could be some argument that $\sqrt{R_{4.123}^2 - R_{4.12}^2}$ is preferable to $R_{4.123} - R_{4.12}$ as a measure of the effect of Block 3 as

distributed, independently of Blocks 1 and 2. These two measures would give quite different estimates. It would be useful to explicate exactly the operational differences (in terms of hypothetical experiments of the sort discussed earlier) between the two measures, but limitations of time, space, and purpose of this paper prevent that here.

Now from the three studies under consideration, we can, with the published data, obtain one measure of the effect of school variables (#3 above) and an approximation for the other (#2 above). The measure $R_{4.123} - R_{4.12}$ or $\beta_{123} - \beta_{12}$ shows the direct effect of school variations on achievement that is distributed independently of home background and type of school, and the approximation to $\beta_{3.21}$, which is a function of R_{123}^2 and R_{12}^2 shows the direct effect of school variations on achievement, including both those that are an indirect consequence of the school working through the home, and those that are independent of it. We will call the first "independent explained variation" and the second "total direct explained variation". The total direct explained variation will always be larger than the independent explained variation, because it includes the latter. The independent explained variation will be larger than added explained variance under some circumstances, smaller under others. Algebra will show that added explained variance exceeds independent explained variation when $R_{123} + R_{12} > 1.0$. Thus when R_{123} and R_{12} are greater than about 0.5, the added explained variance will be greater. The added explained variance of course does not have such a straightforward meaning as independent explained variation or total direct explained variation.

It is useful to see what happens to the countries when the two measures I have proposed are used for measuring the effect of school variables, in place of the measure used by the authors. The tabulation is given below

for Population II, Literature (p. 163). (All measures are multiplied by 100.)

	$R_{123}^2 - R_{12}^2$ Added Explained Variance		$R_{123} - R_{12}$ Indep. Explained Variance		$\frac{1}{2}\sqrt{R_{123}^2 - R_{12}^2}\,(1+\sqrt{1-R_{12}^2})$ Total Direct Expl. Variance	
	value	rank	value	rank	value	rank
Belgium (Flemish-speaking)	6.9	6	6.4	7	28.4	5
Belgium (French-speaking)	9.0	2	8.3	3	32.3	1
Chile	8.9	3	8.9	2	31.7	3
England	3.9	9	3.1	10	22.3	9
Finland	5.5	8	5.0	8	25.5	8
Iran	12.1	1	17.3	1	27.1	6
Italy	3.0	10	3.3	9	18.3	10
New Zealand	8.7	4	7.5	5	32.3	1
Sweden	6.1	7	7.2	6	25.8	7
United States	7.7	5	7.8	4	29.5	4

This tabulation shows that the measure of independent explained variation, $R_{123} - R_{12}$ is close to the added variance measure used by the authors. The second measure, of the total direct explained variation in achievement by school variables, is much larger, and ranks the countries much differently. Now the two measures I have proposed can be compared, because they are both measures of the proporation (or multiplied by 100, percentage) of the difference between any two percentile scores that is accounted for by the same percentiles in the independent variables, considered as a best-predicting composite. The measure of independent explained variation is the effect of these school variables, as distributed, independently of the home backgrounds that partly determine their distribution. The measure of total direct explained variation is the total effect of these school

variables if they were distributed independently of home background.

The first measure may be thought of as the actual effect of the schools in interrupting the transmission of home background resources across generations, or the effect of the schools in equalizing educational opportunity. The second, as the total effect of the schools, is the potential the schools have for doing this if they were in fact distributed independently of home backgrounds. Countries can be compared according to the proportion of their schools' total impact that is distributed independently of home background resources - and, of course, type of school and program, which is not itself distributed wholly in accord with home background. Such a comparison, using the same countries and same age group as before is shown below. Except for Iran,

	Total Explained by Block 3	Proportion Independent of Blocks 1 & 2
Belgium (Flemish)	28.4	.23
Belgium (French)	32.3	.25
Chile	31.7	.28
England	22.3	.14
Finland	25.5	.20
Iran	27.1	.63
Italy	18.3	.18
New Zealand	32.3	.23
Sweden	25.8	.28
United States	29.5	.26

where a very large proportion of the schools' impact is independent of home background (though so little variation is explained by home background in Iran [$R_1 = .21$, much less than in any other country] that there may have been serious measurement problems), only about a quarter of the school variables' influence is independent of home background and type of school.

Among all countries, England is lowest. This accords with qualitative impressions, according to which England's schools are still more strongly class stratified than in most countries. As we will see in subsequent tabulations, England is consistently lowest in this measure of equalizing opportunity through the schools.

Now, finally, it is possible to get on with the task of examining the implications of these findings for our knowledge of effects of the school on learning. In doing this, I will focus on only those six countries that were covered in all three studies: Chile, England, Finland, Italy, Sweden, and the U.S.

II. Substantive Results

Between-school and Between-student Analyses:

In all three studies, there were regression analyses carried out in which the total variation to be accounted for was that between school averages, and regression analyses in which the total variation to be accounted for was the full variation between individual students. Which should be used for assessing the effects of school variables?

The between-schools analysis, by standardizing the variance between schools to 1.0, prevents one question from being asked: what is the overall impact of measured school variables upon achievement, compared to all the other variables that make some individuals achieve more highly than others? Since the strength and distribution of school variables will partly determine the amount of variation among schools, the standardization of the between school variance to 1.0 eliminates a portion of the effect that school variables have. Further, this unknown portion may vary from country to country. Only if one used raw regression coefficients to evaluate the

effects of schools would the measure of effect be comparable in the between-schools and between-students analyses. But because one is interested in the effects of clusters of school variables (partly because any single variable has a very small effect), for which raw coefficients cannot be used, then measures much as those discussed in preceding sections are necessary.

For this reason, in examining overall effects of school variables, the between-student analysis should be used. This analysis has the defect that it includes a lot of variance due to individual differences within schools, which cannot be explained by variables that are constant for a school. Thus the effect of school variables appears especially small. This is one reason between-school analyses are carried out. However, for the reasons discussed above, it appears better to recognize that school variables can account only for that fraction of the total variance that lies between schools, and to use the between-student analysis for the study of effects of school variables.

In use of the between-student analysis, I will focus on the six countries which engaged in testing in literature, reading, and science. These are Chile, England, Finland, Italy, Sweden, and the United States. In addition, I will confine my attention to the 10 and 14 year-olds. Population IV, the last year of secondary school, is such a non-random sample of the population of the age cohort that an examination of the amount of variance or variation accounted for by differences among school resources is not likely to prove very productive. Beyond this, literature was not administered for the 10-year olds. Altogether, I will examine a very partial set of data.

For these six countries, I have calculated the measures described

above, which I will restate here:

1. The total effects of home background R_1
2. The total direct effects of school type and program $\beta_{2.1}$
3. The effects of school type and program distributed independently of home background $R_{12} - R_1$
4. The proportion of school type effects that are distributed independently of home background $(R_{12} - R_1)/\beta_{2.1}$
5. The total direct effects of school variables $\beta_{3.12}$
6. The effects of school variables distributed independently of home background and school type and program $R_{123} - R_{12}$
7. The proportion of school variable effects that are distributed independently of home background and school type and program $R_{123} - R_{12}/\beta_{3.12}$

Of these measures, I will neglect 2, 3, and 4, having to do with school type and program, because the "effects" of school type and program are primarily effects of selection of differently-achieving students into different programs or schools. The school effects, insofar as they exist, are to be found in the specific school variables, while the school type is more a consequence of achievement than a cause.

Table 1 and 2 show measures 1, 5, 6, and 7 for literature, reading, and science for these six countries, for the 14 year-olds, and for reading and science for the 10 year-olds. The first table shows, for example, that in Chile 38% of the variation among 14 year-olds' achievement in literature is accounted for by the total effect of home background, including both direct effects and effects through the schools. In the same country and subject, 9% of the variation is accounted for by school variables that are distributed independently of the home background and school type variables.

Table 1

Population II: age 14

		Chile	England	Finland	Italy	Sweden	U.S.	Average
Total home background effects								
$R_1 = \beta_1$	Literature	.38	.50	.43	.33	.39	.43	.41
	Reading	.45	.52	.45	.32	.40	.47	.44
	Science	.36	.48	.47	.32	.42	.47	.42
	Average	.40	.50	.45	.32	.41	.46	.42
School Effects distributed independently of home & school type								
$R_{123} - R_{12} = \beta_{123} - \beta_{12}$	Literature	.09	.03	.05	.03	.07	.08	.06
	Reading	.06	.02	.04	.03	.04	.06	.04
	Science	.07	.05	.09	.07	.08	.07	.07
	Average	.07	.04	.06	.05	.06	.07	.06
Total direct school effects								
$\beta_{3.12}$	Literature	.32	.22	.26	.18	.26	.30	.26
	Reading	.28	.19	.23	.19	.18	.28	.22
	Science	.26	.30	34	.26	.28	.28	.29
	Average	.28	.25	.28	.21	.24	.29	.26
Proportion of school effects distributed independently of home and school type								
$\frac{R_{123} - R_{12}}{\beta_{3.12}}$	Literature	.28	.14	.20	.18	.28	.26	.22
	Reading	.22	.11	.16	.17	.20	.21	.18
	Science	.26	.18	.25	.29	.27	.24	.25
	$\frac{av(R_{123} - R_{12})}{av\beta_{3.21}}$	.25	.14	.23	.22	.26	.24	.22

Table 2

Population I: age 10

		Chile	England	Finland	Italy	Sweden	U.S.	Average
Total home background effects								
	Reading	.12	.47	.42	.31	.34	.45	.35
	Science	.20	.46	.37	.20	.40	.42	.34
	Average	.16	.46	.40	.26	.37	.44	.35
School effects distributed independently of home & school type								
	Reading	.17	.02	.03	.06	.04	.04	.06
	Science	.16	.03	.05	.08	.06	.09	.08
	Average	.17	.02	.04	.07	.05	.07	.07
Total direct school effects								
	Reading	.29	.13	.18	.22	.18	.21	.20
	Science	.30	.18	.21	.20	.23	.32	.24
	Average	.30	.16	.19	.21	.21	.26	.22
Proportion of school effects distributed independently of home and school type								
$\frac{R_{123} - R_{12}}{\beta_{3.12}}$	Reading	.58	.12	.18	.30	.23	.19	.27
	Science	.49	.18	.24	.41	.25	.29	.31
	$\frac{av(R_{123} - R_{12})}{av(\beta_{3.21})}$	.56	.15	.21	.35	.24	.25	.31

But the total direct effect of the school variables is such they account for 32% of the variation in literature achievement when home background and school type are controlled, so that the 9% constitutes only 28% (9/32) of the total school effect. The remaining portions of the two tables can be read in the same way.

First, several general results are useful to state. For all three subjects, the total effect of home background is considerably greater than the total direct effect of school variables. The overall average is .42 for home background, but only .26 for school at age 14. For population I, the 10 year-olds, school variables are higher relative to home background (.22 to .35 for the overall averages), but because of measurement differences for home background at the two ages, this might not be a true difference. These comparisons show a much higher relative effect of school variables to home background than is ordinarily reported - partly, as I indicated earlier, because of the methods of reporting, which report something like the second set of rows, the effect of school variables distributed independently of home background. As the fourth set of rows shows, this independently distributed effect of school variables is only about 20 to 30 percent of the total, showing that most of the variations in school resources go to reinforce home background rather than to cross-cut it.

Comparisons between ages 10 and 14 show, in addition to the home background differences discussed above, a slightly higher total direct effect of schools at age 14 than at age 10 (.26 to .22 on average) but a smaller proportion of it independent of home background (.22 to .31 on average).

Going beyond these overall comparisons, there are a number of infer-

ences that one can draw from the results shown in these tables. One of the most interesting concerns a comparison between reading, on one hand, and literature on the other. Looking at the averages over countries shows several things. First, home background explains slightly more variation in reading than in either literature or science for the 14 year-olds, or science for the 10 year-olds. But if we examine the total direct effects of school at both age levels, school variables account for somewhat less variation in reading achievement than does literature or science at age 14, or science at age 10. Furthermore, looking at the last set of measures for both 14 year-olds and 10 year-olds shows that a smaller fraction of these (smaller) school effects is distributed independently of home background for reading than for either of the other two subjects.

These data show fairly conclusively, then, that reading achievement is more fully an outgrowth of home influences than are either of the other two subjects, less a function of what takes place at school. This is a rather important result, because it indicates that the general finding in this study and others that home background is a much more powerful influence than school influences in determining achievement is a result that is subject-specific. Not all subjects are alike in the mix of family influences and school influences that determines their achievement.

Beyond these comparisons by age group and by subject matter, there are others that can be made about specific countries. England is perhaps the most consistent, and because of that, the most interesting to look at. First, for both ages, England shows the highest proportion of variation in achievement explained by variations in home background (.50 for 14 year-olds, .46 for 10 year-olds. Secondly, and most interesting, the fourth set of rows shows that England has, in every subject and at both ages, the smallest

proportion of total direct effects of school variables distributed independently of home background. This is the case even though, as the third set of rows in each table shows, England does not have especially high total school effects. The fourth set of rows can be regarded as a measure of the equality of educational opportunity for children of different home backgrounds. The higher the measure, the larger the proportion of variation in achievement due to school resources that are distributed independently of home background resources. On this, England is the only country that is consistently different from (lower than) the others, but Chile shows an exceptionally high figure at age 10.

These are the major inferences I am able to draw from these studies on the effects of school characteristics on learning. The growth scores, the retentivity analyses, the proportion of variance between schools (not reported in Science) did not lead, for me, to further insights about effects of schools on learning. I am sure, however, given the richness of the data and the general sophistication of the analyses to which it has been subject, that there are other important ones, some already reported in these studies, and some which will be the result of further analyses of these data. Altogether, although survey data are sharply limited in their ability to show the effects of school characteristics upon learning, these studies, and the data underlying them, will constitute in the years to come one of the most important sources of insight into the effects of school resources in learning.*

* I have not here examined the effects of any specific school variables, but have left these variables in their original "block". My reason for doing so lies in my belief that survey methods are simply not capable of analyzing such fine grain effects of school variables, and that they are useful only for the more gross questions of the sort I have discussed in this section. I should note that in these studies as in most comparable ones (my own EEOS study included), valiant attempts were made to find effects of specific school variables, but without consistent results except for some variables, such as grade in school, which are merely indicators of the child's general achievement level.

Appendix

Calculation of standardized regression coefficient for a compound of variables, controlling on another set of variables

Let the variables be labelled as follows:

1: a linear combination of Block 1 variables, created as the best-fitting combination for predicting achievement

2: a linear combination of Block 2 variables, the best-fitting combination for predicting achievement

4: achievement

Then $\beta_{42.1}$ is the standardized regression coefficient for Block 2 variables in explaining achievement, with Block 1 variables fixed.

$R_{4.21}$ is the multiple correlation of Blocks 1 and 2 together with achievement.

R_{41} is the multiple correlation of Block 1 variables with achievement.

R_{42} is the multiple correlation of Block 2 variables with achievement.

R_{12} is the correlation of the compound of Block 1 variables with the compound of Block 2 variables.

Then
$$\beta_{42.1} = \frac{\sqrt{R_{4.21}^2 - R_{41}^2}}{\sqrt{1 - R_{12}^2}} \qquad (A.1)$$

$R_{4.21}$, R_{41}, and R_{42} may be obtained directly as the multiple correlations of regressions of achievement on Blocks 1 and 2, Block 1, and Block 2, respectively. R_{12} may be obtained from these three quantities by use of the following equation:

$$R_{12} = \frac{1}{R_{4.21}^2}\left[R_{42}R_{41} - \sqrt{R_{42}^2R_{41}^2 + R_{4.21}^4 - R_{4.21}^2(R_{41}^2 + R_{42}^2)}\right] \qquad (A.2)$$

Having $R_{4.21}$, R_{41}, and R_{42}, one may calculate R_{12} and then $\beta_{42.1}$, or alternatively $\beta_{41.2}$, by use of R_{42}^2 in place of R_{41}^2 in equation (A.1).*

Another method by which $\beta_{42.1}$ may be calculated is to use the regression weights from the regression equation including Blocks 1 and 2 to calculate zero-order correlations of the new compound variable with all other variables. The equation for doing so is given below, where:

r_{ij} = zero-order correlation between variables i and j

r_{cj} = zero-order correlation between variable j and the new compound variable

σ_i = standard deviation of variable i

$b_{i.s}$ = multiple regression coefficient of variable i with the dependent variable in the presence of variables 1,...,n, the set of variables to be labelled s

Variables 1,2,...,n are not to be compounded, and variables n+1,...m are to be compounded.

Then

$$r_{cj} = \frac{\sum_{i=n+1}^{m} b_{i.s} r_{ij} \sigma_i \sigma_j}{\sigma_j \left[\sum_{i=n+1}^{m} b_{i.s}^2 \sigma_i^2 + 2 \sum_{k=i+1}^{m} \sum_{i=n+1}^{m} b_{i.s} b_{k.s} r_{ik} \sigma_i \sigma_k \right]^{\frac{1}{2}}} \qquad \text{(A.3)}$$

* It should be noted that this method for calculating $\beta_{42.1}$, in contrast to the one given below, involves an approximation. Its virtue lies in the fact that it may be calculated by hand after regression analyses have been done, assuming that $R_{4.21}$, R_{41}, and R_{42} have been obtained in the regression analyses. The approximation lies in the fact that the method assumes that "variable 1", and "variable 2" appearing in the equation that gives $R_{4.21}$ are exactly the same variables which appear in the equations that give R_{41} and R_{42}. If the regression weights are the same relative size for variable 1 in both equations in which it enters, and for variable 2 in both equations in which it enters, then the compounds will be the same, and the assumpation holds. If not, then a slight error is introduced, because the compound that gives R_{41} is designed to maximize it, thus producing a slightly higher value of R_{41} than the compound designed to maximize $R_{4.21}$ would do.

The same calculation may be used for the correlation of the compound variable with the dependent variable. Once these new zero-order correlations have been calculated, then the new regression, and the desired standardized regression coefficient for the new Block 2 compound, may be calculated.

These two methods of calculating standardized regression coefficients for new compounds may of course be incorporated into a computer program so that calculations for the compound are automatically done by specifying in advance the compounds for which standardized coefficients are desired.

When only $R_{4.21}$ and R_{41} have been presented, as in the studies under consideration, then some reasonable bounds for $\beta_{42.1}$ may be calculated as follows:

The minimum of R_{12}^2 is 0, so that the maximum of the denominator of A.1 is 1.0, and thus the minimum of $\beta_{42.1}$ is $\sqrt{R_{4.21}^2 - R_{41}^2}$. As a reasonable, though not definite, upper bound, we note that R_{41} is ordinarily greater than R_{12}, since the variables 1 and 2 have been selected to correlate with 4, and not with each other. Thus ordinarily, $\sqrt{1 - R_{41}^2}$ will be smaller than $\sqrt{1 - R_{21}^2}$. Consequently, a reasonable upper bound for $\beta_{42.1}$ will be $\sqrt{R_{4.21}^2 - R_{41}^2} / \sqrt{1 - R_{41}^2}$. Thus the following inequalities can be used to obtain an estimate for $\beta_{42.1}$:

$$\frac{\sqrt{R_{4.21}^2 - R_{41}^2}}{\sqrt{1 - R_{41}^2}} > \beta_{42.1} > \sqrt{R_{4.21}^2 - R_{41}^2}$$

One might then estimate $\beta_{42.1}$ by averaging these two bounds. For the case of England in Literature given earlier (p. 163), the estimates are:

$$R_{41}^2 = .252$$

$$R_{4.21}^2 = .371$$

$$R_{4.321}^2 = .401$$

$$\sqrt{\frac{.371 - .252}{1 - .252}} = \sqrt{\frac{.119}{.748}} > \beta_{42.1} > \sqrt{.119}$$

$$.399 > \beta_{42.1} > .345$$

$$\beta_{42.1} \approx .372$$

$$\sqrt{\frac{.410 - .371}{1 - .371}} = \sqrt{\frac{.039}{.629}} > \beta_{43.21} > \sqrt{.039}$$

$$.249 > \beta_{43.21} > .197$$

$$\beta_{43.21} \approx .223$$

These bounds are not extremely wide, and thus provide some useful information.